UofT - CIPO Database Structure

abstract table

The table includes a single paragraph describing the invention written in both English, text_en, and French, text_fr. It also includes the doc_id, which corresponds to the [document number](#) of the patent. Document numbers increase over time, where newer patents will have higher document numbers.

agent table

The table includes the agent's name (the patent company that represented the owner to the patent office), as well as the doc_id, which corresponds to the [document number](#) of the patent. A patent document has a single agent.

applicant table

The table includes the applicant's name (this can be a person or a corporate entity), address, city admin_div (such as province or state), country, nationality and residence. Not all values are present in all cases. It also includes the doc_id, which corresponds to the [document number](#) of the patent, and a sequence number, which corresponds to the order in which the name appeared in the patent. A patent document may have many applicants.

assignee table

The table includes the assignee's name (this can be a person or a corporate entity), address, city admin_div (such as province or state) and country. Not all values are present in all cases. It also includes an enable date field, indicating the date on which the owner received all or part of the ownership, and a disable date field, indicating the date on which the owner stopped having ownership. A grantee boolean field indicates if the owner is the grantee of the patent where t = True/Yes and f = False/No. The doc_id corresponds to the [document number](#) of the patent. A patent document may have many assignees or owners, and not all may be active.

claim table

The table contains the doc_id, which corresponds to the [document number](#) of the patent, and the text of the original patent disclosure statement. This text is stored as raw XML, and contains html and other markup language. The fields in the table should not be accessed by print to screen, as the large amount of raw text inside of them can break the display. Please use

PostgreSQL XML functions to access these data, and save any results as XML documents. See sample queries for more information on querying these data.

disclosure table

The table contains the doc_id, which corresponds to the [document number](#) of the patent, and the complete description text of the patent. This text is stored as raw XML, and contains html and other markup tags. The fields in the table should not be accessed by print to screen, as the large amount of raw text inside of them can break the display. Please use PostgreSQL XML functions to access these data, and save any results as XML documents. See sample queries for more information on querying these data. Please note that this field may be blank in some cases, particularly for older patent documents.

document table

The table contains high-level metadata about the patent document, where id corresponds to the [document number](#) of the patent and date corresponds to the BBLO (publication date) of a published patent, or AAFI (filing date) for a non-published patent; see **history table** below for more information on dates. The table also includes values for document kind; see WIPO's [kind code element guide](#) for descriptions. A boolean pct field indicates if the application was filed through the [Patent Cooperation Treaty](#)(PCT) where t= PCT compliant and f = Non-PCT compliant. app_country lists the original filing country of the patent application, where app_number is the document number of that original application, app_date the original filing date, and filing_lang the original filing language (note that if the original filing country was Canada, the value of app_number will match id). updated contains the last update date of the application or registration.

history table

The table includes a date as well as a status, which indicates the event represented by that date. It also includes the doc_id, which corresponds to the [document number](#) of the patent. Most doc_ids will have many date and status entries, for example: AAFI is the date the patent was first filled; BBLO is the date the patent was opened to public inspection; FGGR is the date the patent was formally issued. Not all date types will be present in all cases. A full list of date status codes is available via the Government of Canada's [Patent Status Dictionary.](#)

inventor table

The table includes the inventor's name (this can be a person or a corporate entity), address, city, admin_div (such as province or state), and country. Not all values are present in all cases. It also includes the doc_id, which corresponds to the [document number](#) of the patent, and a sequence, which corresponds to the order in which the name appeared in the patent. A single patent document may have many inventors.

icpr table

The table contains data about the patent document related to the [International Patent Classification (IPC) code](#). Detailed information on possible values for the fields listed below can be found in the IPC code documentation. level contains the classification level indicator for the IPC; section contains the Section Symbol of the IPC, designated by one of the capital letters A through I; class contains the Class Symbol, where Classes are a subdivision of Sections and each Class has a Class Title (Ex: G06 Computing); subclasses contain the Subclass Symbol, which are a further subdivision of classes. main_group and subgroup are further subdivisions of Subclasses. symbol_position contains the position of the first invention information classification, and can be F (first) or L (later), while value represents the classification value indicator for I (invention) or N (non-invention). action_date contains the date on which the classification information was assigned, and generating_office indicates the country who assigned the information. status represents the type of data, date_source contains the identifier of the source of the classification data, and version contains the version of the International Patent Classification (IPC) used. The table also includes the doc_id, which corresponds to the [document number](#) of the patent.

parent_document table

The table contains information on the parent document of a patent document. Many component patents may be grouped under a single parent patent, or a patent may represent an improvement to an original parent patent. It includes the parent_app_number, or document number of the parent patent, and a corresponding divisional date. It also includes the doc_id, which corresponds to the [document number](#) of the patent.

priority_claim table

Canada is a signatory to international treaties and conventions which grant similar rights throughout many countries. A patent applicant may use the filing date of a previously filed application as the effective filing date of subject matter in a later filed Canadian application. The table provides information on those priority application(s) related to the Canadian application. It contains a kind field indicating if the priority claim is from a foreign country i.e. "international", the two-letter country code of the priority application, the foreign_app_num or document number of this application, the date of the corresponding priority application, and a sequence number, which corresponds to the order in which the name appeared in the patent. It also includes the doc_id, which corresponds to the [document number](#) of the patent. A patent document may have multiple priority claims.

title table

The table includes the name of the invention as provided by the applicant, in both English, text_en, and French, text_fr. It also includes the doc_id, which corresponds to the [document number](#) of the patent.